Defining white matter: towards standardizing nomenclature, ontology, and taxonomy of white matter pathways, and practices to dissect them with diffusion MRI fiber tractography

Thank you for your interest in this multi-institutional collaborative project! If you have any questions at any time, please do not hesitate to contact Kurt Schilling (kurt.g.schilling.1@vumc.org), or Maxime Descoteaux (maxime.descoteaux@usherbrooke.ca).

# Overview:

The aims of this project are three-fold: (1) provide a reference for “how” the field tracks major white matter pathways using state of the art methods, (2) spur discussion on how we “define" the white matter pathways, and how this definition leads to constraints and procedures in the tractography process, and (3) release this as both dataset (data, streamlines, volumes) and code-based resource, freely sourced, to neuroimaging community.

We have performed a targeted review of neuroimaging literature and will plan on extending >30 invites for collaboration. As a collaborator, you will be asked to extract (i.e., virtually extract) up to 14 white matter pathways from six diffusion MRI fiber tractography datasets. This is NOT a challenge, nor a competition, but rather a collaborative effort to better understand the taxonomy, nomenclature, and ontology of the human brains white matter, and improve upon our efforts to study this white matter using tractography. All collaborators will be included as co-authors (if desired and indicated in the “collaboration agreement”) in all manuscripts or written communications that result from this project.

# Tasks:

The task is to extract 14 major white matter pathways on the six diffusion MRI datasets provided. We have already run tractography and provide the streamlines for each subject (see Provided Data, below). We provide whole-brain tractography using both deterministic tractography and probabilistic tractography methods, and you are free to choose whichever (or both) are best suited for your virtual dissection approach.

We ask that you extract the LEFT pathway for each of the pathways below (we note that several are commissural pathways and streamlines from the RIGHT hemisphere should be included if the pathway does indeed enter the RIGHT hemisphere). You are free to utilize any software that you desire (we recommend Trackvis or MI-brain). We will provide guidance with these softwares, and any assistance you may require regarding file formats, processing, submissions, etc.

Thus, there are two primary tasks.

**(1) Define**the white matter tracts. A definition might (or might not) describe where it starts, where it ends, where it goes, its relation to cortical structure, its shape, and its location relative to other structures. A definition might (or might not) even include the type of pathway, if it is part of a network of pathways, if it has sub-divisions, or even it is exists!

**(2) Dissect**the streamlines that make up these pathways. And importantly, describe how your definition influences your dissection process, your constraints, or your decisions.

Remember to save everything. Atlases used, inclusion/exclusion regions, scene files, TRK/TCK files, code (exact code and versions), etc. It is critical that every process is reproducible (within reason) given your description and/or figures!

Important information

* Streamlines are created by seeding throughout all white matter, and end once they enter gray matter.
* If pathways have sub-divisions (or are part of a hierarchy of a larger bundles or system) please note and describe this (see Written Description, below)
* If the existence of the pathway is contentious, or in doubt, we ask that you describe this (you do not have to track these pathways if you believe it does not exist)
* You do not have to define all pathways. If you do not have a definition or protocol, or if you are uncertain about a given pathway, we allow partial submissions (i.e. less than 14 tractograms).
* You are also free to define more than the requested pathways. While we may not explicitly detail them in the first series of studies, we envision this leading to future studies covering a range of connections throughout the brain.
* Also note that if differences in protocol/definition would exist for the RIGHT hemisphere, please describe this (you are also free to track the right hemisphere on these subjects if desired, but this is not required)
* It is possible some pathways are “hard to track” and may not exist in the data, please note this
* Data has been processed using the HCP pre-processing pipeline, and is oriented in MNI space (tractography is performed in this space).
* We do not encourage this, but you are additionally free to create your own streamlines using your chosen approach (on the same six subjects). This will make direct comparisons of streamlines impossible, but volumes can still be compared. If this is done, you MUST perform tracking in diffusion space. If you intend to do this, we will provide the raw diffusion data as requested.
* If you need additional data, please let us know.

Pathways to reconstruct

1. Superior Longitudinal Fasciculus (SLF)
2. Arcuate Fasiculus (AF)
3. Optic Radiations (OR)
4. Corticospinal Tract (CST)
5. Cingulum (CG)
6. Uncinate fasciulus (UF)
7. Corpus Callosum (CC)
8. Middle Longitudinal Fasiculus (MLF)
9. Inferior Fronto-occipital Fasiculus (IFOF)
10. Inferior longitudinal Fasiculus (ILF)
11. Fornix (FX)
12. Anterior Commisure (AC)
13. Posterior Commisure (PC)
14. Parieto-occopital pontine tract (POPT)

# Provided data

The data repository is within Vanderbilt Box: <https://vumc.box.com/s/j33mgl8kfrjcvfl1d9wzf75gdlf1ek6j>. Within this repository, you will find a “Data” folder that contains the data to be downloaded, as well as a “submissions” folder where results will be placed. Please download the latest version of this folder.

Within the data folder you will find 6 folders, named “s1” through “s6” designating the six unique subjects.

Within each subject folder, you will find (all data is in the same space, MNI aligned, 1mm isotropic):

B0.nii.gz: non-diffusion weighted image

Fa.nii.gz: FA map

Mask-brain.nii.gz: brain mask

Mask-csf.nii.gz: CSF mask

Mask-gm.nii.gz: gray matter mask

Mask-wm.nii.gz: white matter mask

RGB-3d.nii.gz: RGB image (each voxel contains a tuple of 3 uint8 values)

RGB-4d.nii.gz: RGB image (4D array with last dimension = 3, set in uint8 and [0,255])

Rgb.nii.gz: RGB image (4D array with last dimension = 3, set in uint8 and [0,1])

T1.nii.gz: T1 image

Tracking-deterministic.trk: deterministic tracking streamlines

Tracking-probabilistic.trk: probabilistic streamlines

* each streamline files are created from whole brain tractography and contain approximately 1.6-1.9 million streamlines
* TRK files have been compressed, if you require uncompressed, we will provide this for you
* We have chosen to provide TRK files (compressed). We have also done conversion to TCK files, but are willing to convert and send alternative file types (FIB)
* note that because of the large number of streamlines, many software packages visualize only subsets of, it is crucial that you make sure you are viewing all streamlines when determining final outputs.

Additionally, to accommodate software and systems with limited capabilities, we have also included the tractograms split into left hemisphere, right hemisphere, and commissural fibers as:

Tracking-deterministic-right.trk: deterministic fibers entirely in right hemisphere

Tracking-deterministic-left.trk: deterministic fibers entirely in left hemisphere

Tracking-deterministic-comm.trk: deterministic fibers with an endpoint in both hemispheres

Tracking-probabilistic-right.trk: probabilistic fibers entirely in right hemisphere

Tracking- probabilistic-left.trk: probabilistic fibers entirely in left hemisphere

Tracking- probabilistic-comm.trk: probabilistic fibers with an endpoint in both hemispheres

# How to get the data

You must fill out a collaboration form to indicate intent to collaborate: <https://docs.google.com/forms/d/e/1FAIpQLSdnYGwYIffKw0fSrCL2DeLAcee6k-gKP-GOCYZxiIzZ_fDUcA/viewform?usp=sf_link>

# Data to submit

We ask for a number of results to be submitted, including data (results), written descriptions, and figures (optional). Note that .doc files can also be submitted as .txt, .docx, etc.

Data:

1. For each pathway (for each subject) save and submit the streamlines as TRK file, for example **XXX.trk** where XXX is the acronym of each pathway (*required*).

2. If used, save any and all constraints and regions used in tracking (AND and NOT regions, etc.) (*required*) - we suggest a MISCELLANEOUS folder, that includes code/scripts, all ROIs used, query files, atlases used for analysis, scene files, and anything that can be used to exactly reproduce the results. For example, if you used TractQuerier, please provide the query file, etc.

Written description:

1. **Methodology.doc** (*required*) (<400 words): Concise description of overall methodology to extract pathways (this does not include pathway specific descriptions, but only a high-level summary of methodology – i.e., clustering, manual dissections, etc).

2. **PathwayDescriptions.doc**: (*required*) (<200 words per pathway): For each pathway, explicitly “define “the pathway (for example: “from Noggle et al., 2011, the Arcuate Fasciculus is composed of fiber bundles that extend anteriorly from the posterior portion of the temporal lobe to the posterior region of the inferior prefrontal lobe, thereby linking the expressive and receptive language centers of the cortex”), and how this definition influences constraints in tracking (be specific with regions of interest, how they were defined, if gray/white matter atlases were used, etc.).

3. **Discussion.doc** (*optional*) (no word limit): this can be any discussion on nomenclature, taxonomy, ontology, and choices made in tracking. As well as any discussion on white matter tractography, improvements needed, limitations of the current project, etc. Depending on subject matter, this content may be used in discussions for manuscripts (upon collaborator’s approval)

4. **Info.doc**: (*required*) document that details names and affiliations of all participating collaborators.

5. **CODE.doc** (*optional*): Any and all code, if used (explicitly detailing software and commands utilized and inputs/outputs). This will be provided as supplementary information.

6. You are welcome to include any additional supplementary information, code, figures of your choosing.

Figures (all figures are optional)

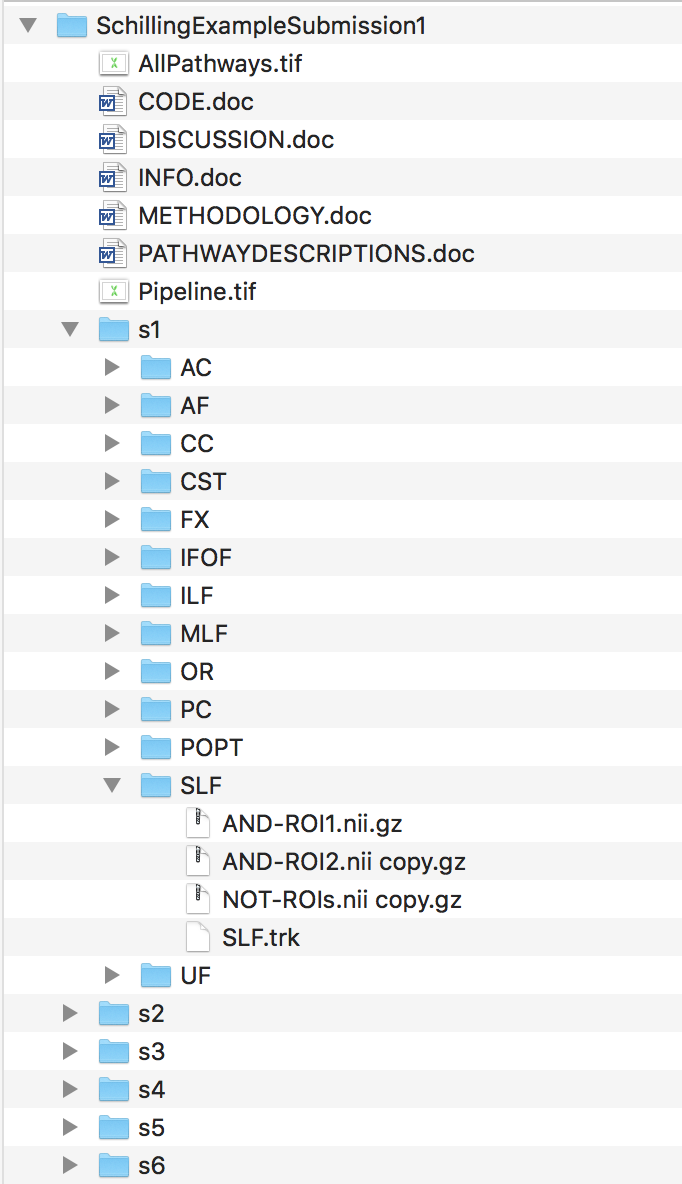
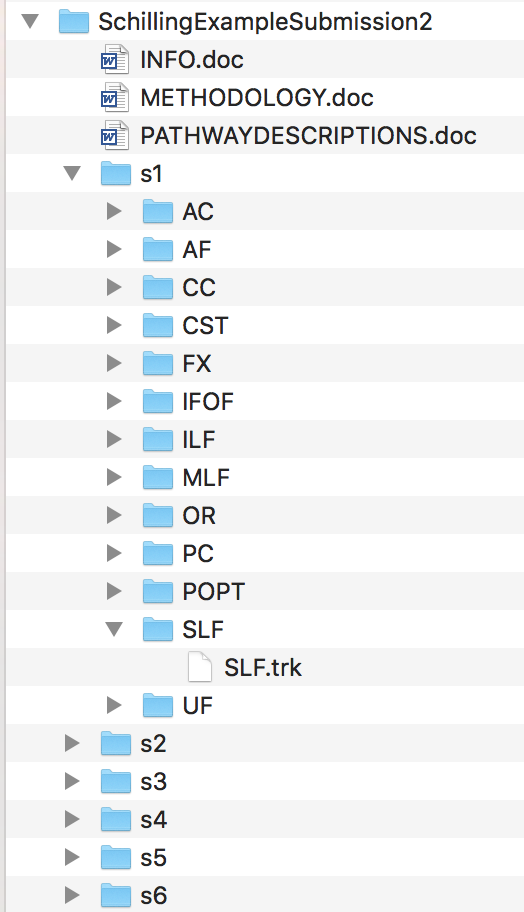
We highly recommend submitting figures that will be utilized in manuscripts. These figures can include summary pipelines (diagrams/cartoons of processing), detailed examples (landmarks, ROIs, clusters, labels, for all pathways), and display of all pathways (for example all pathways of a given subject) displayed however you would like. These are not required, but will facilitate methodology and discussions. We highly encourage figures that detail the processing pipeline as well as detailed examples of pathway dissection. Do not hesitate to include “more” than necessary, as we envision resulting manuscripts will have supplementary figures and details.

# How to Submit

Submissions will be made in Vanderbilt Box. When ready to submit, email the research leaders (Kurt Schilling and Maxime Descoteaux) to request a submission link. A unique link will be provided for each submission. Please drop your “submission” folder into the linked deposit box, with a unique submission name (note that the name will be hashed/anonymized when all submissions are aggregated into one location).

Your created folder shall contain several .doc (or .docx, .txt, etc) files, as well as optional figures, and 6 subject folders (named s1-6). Each subject folder contains one sub-folder per white matter pathway (named after the acronym – i.e., “SLF”) and each of these contains (at a minimum) one .trk file, named again using the acronym (i.e., “SLF.trk”).

An example submission, containing the bare-minimum requirements, is shown below (left):



Note that the minimum contains written information: info.doc, methodology.doc, pathwaydescriptions.doc; as well as streamlines (as a .trk file) for each subject and each pathway. Each pathway should contain a .trk file (not shown in figure for clarity).

An example submission with additional information, including AND and NOT ROIs, figures, code.doc, discussion.doc, is shown at right. Note these additional files are optional.

# Timeline

All data and results should be submitted by 4 December, 2019. Please inform research leaders (Kurt or Maxime) if more time is required.